BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 2: e25260
doi: 10.3897/biss.2.25260

OPEN ACCESS

Conference Abstract

A Methodological Proposal for Collecting and Creating Macroscopic Photograph Collections of Tropical Woods with Potential for Use in Deep Learning

Erick Mata-Montero‡, Juan Carlos Valverde‡, Dagoberto Arias-Aguilar‡, Geovanni Figueroa-Mata‡

‡ Costa Rica Institute of Technology, Cartago, Costa Rica

Corresponding author: Erick Mata-Montero (erick_mata@yahoo.com)

Received: 25 Mar 2018 | Published: 25 Mar 2018

Citation: Mata-Montero E, Valverde J, Arias-Aguilar D, Figueroa-Mata G (2018) A Methodological Proposal for Collecting and Creating Macroscopic Photograph Collections of Tropical Woods with Potential for Use in Deep Learning. Biodiversity Information Science and Standards 2: e25260. https://doi.org/10.3897/biss.2.25260

Abstract

Costa Rica is one of the countries with highest species biodiversity density in the world. More than 2,000 tree species have already been identified, many of which are used in the building, furniture, and packaging industries (Grayum et al. 2003). This rich diversity makes the correct identification of tree species very difficult. As a result, it is common to see in the national market that species are commercialized with mistaken identifications, which makes quality control particularly challenging. In addition, because 90 timber tree species have been classified as “threatened” in Costa Rica, correct identifications are indispensable for law-enforcement.

The traditional system for tree species identification is based on macro and microscopic evaluations of the anatomy of the wood. It entails assesing anatomical features such as patterns of vessels, parenchymas, and fibers. Typically, 7.7 x 10 cm pieces of wood cuts are used to identify the tree species (Pan and Kudo 2011, Yusof et al. 2013). However, assessing these features is extremely difficult for taxonomists because properties of the wood can vary considerably due to environmental conditions and intra-specific genetic variability.

© Mata-Montero E et al. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Deep learning techniques have recently been used to identify plant species (Carranza-Rojas et al. 2017a, Carranza-Rojas et al. 2017b) and are potentially useful to detect subtle differences in patterns of vessels, parenchyma, and other anatomical features of wood. However, it is necessary to have a large collection of macroscopic photographs of individuals from various parts of the country (Pan and Kudo 2011). As a first step in the application of deep learning techniques, we have defined a formal, standard protocol for collecting wood samples, physically processing them, taking pictures, performing data augmentation, and using metadata to provide the primary data necessary for deep learning applications. Unlike traditional xylotheque sampling methods that destroy trees or use wood from fallen trees, we propose a method that extracts small size samples with sufficient quality for anatomical characterization but does not affect the growth and survival of the individual.

This study has been developed in three forest permanent plots in Costa Rica, all of which are sites with historical growth data over the last 20 years. We have so far evaluated 40 species (10 individuals per species) with diameters greater than 20 cm. From each individual, a cylindrical sample of 12 mm diameter and 7.5 cm in length was extracted with a cordless drill. Each sample is then cut into five of 8 x 8 x 8 mm cubes and further processed to result in curated xylotheque samples, a dataset with all relevant metadata and original images, and a dataset with images obtained by performing data augmentation on the original images.

Keywords

Deep Learning, Automated tree species identification

Presenting author

Erick Mata-Montero

Presented at

Biodiversity Information Standards (TDWG) 2018, Dunedin, NZ

Funding program

Proyecto 1370004, Vicerectoría de Investigación y Extensión, TEC

References

- Carranza-Rojas J, Goeau H, Bonnet P, Mata-Montero E, Joly A (2017a) Going deeper in the automated identification of Herbarium specimens. BMC Evolutionary Biology 17 (1): 1-14. https://doi.org/10.1186/s12862-017-1014-z
- Carranza-Rojas J, Joly A, Bonnet P, Goëau H, Mata-Montero E (2017b) Automated Herbarium Specimen Identification using Deep Learning. Proceedings of TDWG 1: e20302. https://doi.org/10.3897/tdwgproceedings.1.20302
- Grayum MH, Hammel BE, Herrera C, Villalobos NZ (2003) Manual de plantas de Costa Rica /B.E. Hammel ... [et al.] editores ; Silvia Troyo, ilustraciones. Monogr. Syst. Bot. Missouri Bot. https://doi.org/10.5962/bhl.title.891
- Pan S, Kudo M (2011) Segmentation of pores in wood microscopic images based on mathematical morphology with a variable structuring element. Computers and Electronics in Agriculture 75 (2): 250-260. https://doi.org/10.1016/j.compag.2010.11.010
- Yusof R, Khalid M, M. Khairuddin AS (2013) Application of kernel-genetic algorithm as nonlinear feature selection in tropical wood species recognition system. Computers and Electronics in Agriculture 93: 68-77. https://doi.org/10.1016/j.compag.2013.01.007